DECLARATION OF EUGENE LEE ISO GOOGLE LLC'S RESPONSE TO THE COURT'S 10/27/22 ORDER TO SHOW CAUSE (DKT. 784)

Redacted Version of Document Sought to be Sealed

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA OAKLAND DIVISION

CHASOM BROWN, et al., individually and
on behalf of all similarly situated,

  Plaintiffs,

  vs.

GOOGLE LLC,

  Defendant.

Case No. 4:20-cv-03664-YGR-SVK

DECLARATION OF EUGENE LEE

1. I am a Google engineer and a team lead for Google's Conflux and Evenflow teams, which are responsible for the logs processing backend for Ads at Google. In this capacity, I regularly work with ██ logs and ██ logs in the course of my day-to-day responsibilities. I have been a Google employee for more than 10 years. I make this declaration of my own personal, firsthand knowledge, and if called as a witness, I could and would testify competently thereto.

██ *Logs*

2. I have reviewed the June 14, 2022 Declaration of Martin Šrámek served by Google in this litigation, which identified, among other logs, the following ██ logs that contain the "maybe_chrome_incognito" bit:

- ████████
- █████████
- ██████████
- ██████████
- ████████

3. The ██ logs" identified in paragraph 2 are logs that contain information that is taken from two or more other logs. As explained in paragraph 1 above, I have experience with these logs because the Conflux and Evenflow teams that I lead are responsible for generating these logs.

1   As explained further below, these ███ logs" do not join authenticated and unauthenticated

2   information.

3   4.   ███

4   ██████   ██████   ██████ ██ ████████████ ██████████ .   Log ████

5   ████████████████████████ contains the results of joining of the following logs that

6   contain only unauthenticated data: ████████████ and ██ of ███ logs: ████████████,

7   ████████████████████, or ████████████████ .[1] These logs are joined in

8   order for ads systems with legitimate need to access fields from either input by reading a single

9   record. For example, ads systems may need to use these logs to compute bid landscapes, which are

10 used to provide guidance for advertisers' bidding strategies by showing how clicks and impressions

11 can change in response to changes in bidding amounts. The maybe_chrome_incognito bit in this log

12 comes from the input logs █████████████, ██████████████████████, or

13 ██████████████████ . Therefore, this ████ log joins only unauthenticated data to other

14 unauthenticated data.

15   5.   ███████████████████████████████,   ████████████

16 ████████████████████████,   and   ████████████████

17 ████████████████████ are " ██████ " logs that contain records keyed to

18 authenticated identifiers.

19   6.   █████████████████████████████ contains the results of joining of

20 the following ████ logs: ████████████████ and ██ of ██ logs: ████████████

21 ████████ or ████████████████████ . These logs are joined in order for ads

22 systems with legitimate need to access fields from either input by reading a single record. For

23 example, ads systems may need to use these logs to compute bid landscapes, which are used to

24 provide guidance for advertisers' bidding strategies by showing how clicks and impressions can

25

26 _____

27 [1]   Upon investigation, I understand that the following statement in the June 14, 2022
    Declaration of Martin Šrámek is inaccurate: "The ████ log ████

28 ████████████████████ contains information regarding user browsing that was
    included in two other logs

change in response to changes in bidding amounts. The maybe_chrome_incognito bit in this log comes from input logs ███████████████ or ███████████ ████████ .

7. ████████████████████ contains the results of joining the following ██████ logs: ████████████████ and ██ of ███ logs: ████████████████, ████████████████, or ████████████████ . These logs are joined in order for ads systems with legitimate need to access fields from either input by reading a single record. For example, these logs may be used to generate models for ads auctions. The maybe_chrome_incognito bit in this log comes from the input logs ███████████████, ███████████████ ████████████, or ████████████████████ .

8. ████████████████████ contains the results of joining the following ██████ logs: ████████████████ and ██ of ███ logs: ████████████████, ████████████████, or ████████████████ . These logs are joined in order for ads systems with legitimate need to access fields from either input by reading a single record. For example, these logs may be used to generate models for ads auctions. The maybe_chrome_incognito bit in this log comes from input logs ████████████████, ████████████████████, or ████████████████ .

9. None of the ██████ logs identified in paragraphs 6-8 above join data from unauthenticated logs with data from authenticated logs.

10. ████████████████████ contains data from personal logs as well as data from unauthenticated logs, but it does not join personal logs with unauthenticated logs. This log contains the results of (i) sorting data from three unauthenticated logs (████████ ███████████, ████████████████, and ████████████████████); and (ii) sorting data from three authenticated logs (████████████████, ████████████ ████████████, and ████████████████████).

11. Attached hereto as **Exhibit A** is a document titled ████████

1 ██████████████████ which I am familiar with because it was prepared by my team at

2 Google. Exhibit A explains that ████████████████████████ (i) is a "special

3 version of ██████████████ + ██████████████████ logs joinable to the

4 ██████████████ log," Ex. A at 1; (ii) contains information from the three authenticated and

5 three unauthenticated logs identified in paragraph 2 above, *id.*; (iii) "supports an efficient way to be

6 joined with ██████████████ logs in order to get the ████████ information together with the

7 query state information," *id.* at 2; and (iv) is "[o]nly useful for joining with

8 ██████████████," *id.* at 1. In simpler terms, ████████████████████ creates a

9 ████ log (instead of ██ separate logs) that can be joined with the ████████████████ log

10 efficiently. Unauthenticated data in ████████████████████ is joined only with

11 unauthenticated data in ██████████████████; and authenticated data in ████

12 ██████████████ is joined only with authenticated data in ██████████████.

13 Authenticated and unauthenticated data in those logs is not joined, and is not linked together to

14 identify the unauthenticated data with a specific user.

15   12.   Exhibit A also explains that ████████████████████████████ "has

16 exactly the same events as the combined ████████████████ and ████████████████

17 logs." Similarly, any given ████ log does not record any data in addition to what is recorded in the

18 underlying logs that the ████ log combines. Thus, any information recorded about IP Addresses,

19 User Agents, cookies, and the maybe_chrome_incognito bit in the ████ logs identified above is a

20 duplicate of the data recorded in the logs from which the ████ logs are created.

21   13.   Before being contacted by counsel for purposes of this declaration, I was not aware

22 of the "maybe_chrome_incognito" bit or any other bit that purports to estimate Chrome Incognito

23 status. In fact, the "maybe_chrome_incognito" bit is recorded in the ████ logs identified in

24 paragraph 2 above only because the ████ logs take wholesale the fields from the input logs that

25 contain this field. As such, when the maybe_chrome_incognito bit was added to the input logs, it

26 was copied to each of the ████ logs identified in paragraph 2 above.

27   14.   To my knowledge, the ████ logs identified in paragraph 2 above have not been used

28 to perform analysis specifically directed at Incognito browsing.

4

15.   I am not aware of any logs at Google that join records of authenticated information with records of unauthenticated information, nor have I ever been aware of any such logs at any time during the course of my employment at Google.

████ *Logs*

16.   The June 14, 2022 declaration of Martin Šrámek also identified, among other logs, the following ██████ logs that contain the "maybe_chrome_incognito" bit:

- ████████████████
- ████████████████████
- ██████████████████
- ████████████████████
- ██████████████████
- ██████████████████████████
- ████████████████████████████
- ██████████████████████
- ██████████████████████
- ████████████████████████████

17.   The ██ ████████████ logs identified in paragraph 16 above join information from input logs (████████████, ████████████████████████, ██████ ████████████, ██████████████████, ██████████████████████, ██ ████████████████████, and ██████████████████████████) with information indicating that an ad was flagged as spam. The ██ ██████████████ logs identified in paragraph 16 above further annotates the ██████████ logs. As explained in paragraph 1 above, I have experience with these logs because the Conflux and Evenflow teams that I lead are responsible for these logging pipelines' backend architecture.

18.   Attached hereto as Exhibit B is a document titled "A Gentle Introduction to Evenflow Spam Pipelines" that was prepared by my team. Exhibit B explains that ████████████████ ██ These ████ are the "information indicating that an ad was flagged as spam" referenced in paragraph 17 above. ████████

1   are generated by Google's spam team and written to ▮▮ different logs: (i) "▮▮▮ ▮▮▮ are written

2   to the ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ log, where ▮▮ is the kind of the events (e.g.

3   Conversions , Interactions , etc)"; and (ii) "Unspam ▮▮▮ are written to the ▮▮▮▮▮▮

4   ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ log." Ex. B at 5. ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮

5   ▮▮▮

6   ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ *Id.* Exhibit B notes that

7   ▮▮▮

8   ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮ *Id.* However, this does not mean that personal

9   (*i.e.*, authenticated) and non-personal (*i.e.*, unauthenticated) information is joined via these

10 pipelines, as the output logs are still separated into personal and non-personal logs, which is reflected

11 in the log names in paragraph 15 above. *See also id.* at 6 ▮▮▮▮▮▮▮▮▮▮▮▮▮

12 ▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮. Additionally, Exhibit B notes that,

13 for the ▮▮▮▮ annotation pipeline, "▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮▮

14 ▮▮▮▮" *Id.* at 7. These logjoining and annotation pipelines "run[] the following logic: 1. It

15 dedupes the input by ▮▮▮▮▮▮▮▮ , which was set by the spam logjoining pipelines by

16 copying from the original ▮▮ ▮▮▮▮▮. It looks up the dependent annotated ▮▮▮▮. It

17 does some ▮▮▮▮ in order to make the month-end decision ('Should we consider this event in

18 this month's bills for advertisers?')[; and] 4. It merges the new month-end decision with the looked-

19 up event and writes the result to the annotated output log." *Id*

20   19.   The maybe_chrome_incognito bit is recorded in the ▮▮▮▮ logs identified in

21 paragraph 16 only because: (1) the ▮ ▮▮▮▮▮▮ logs take wholesale the fields from the

22 input logs that contain these fields: ▮▮▮▮▮▮▮, ▮▮▮▮▮▮▮▮▮▮▮▮▮,

23 ▮▮▮▮▮▮▮▮▮, ▮▮▮▮▮▮▮▮▮, ▮▮▮▮▮▮▮▮▮,

24 ▮▮▮▮▮▮▮▮▮▮, and ▮▮▮▮▮▮▮▮▮▮; and (2) the

25 ▮ ▮▮▮▮▮▮ logs take wholesale the fields from the ▮▮▮▮▮ logs. As such,

26 when the maybe_chrome_incognito bit was added to the input logs, it was copied to each of the
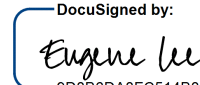
27 ▮▮▮ logs identified in paragraph 16 above.

28

1   20.   ███████ logs do not record any further action from the user (*e.g.*, views, interactions,

2 or clicks). Thus, the only event-level browsing data in these ███████ logs (including the

3 "maybe_chrome_incognito" field) is duplicated directly from the input logs identified in paragraph

4 19.

5   21.   To my knowledge, the ███████ logs identified in paragraph 16 above have not been

6 used to perform analysis specifically directed at Incognito browsing.

7

8   I declare under penalty of perjury that the foregoing is true and correct.

9   Executed on the 30th day of November, 2022 at Kirkland, Washington.

10

11   DocuSigned by:

    By: *Eugene Lee*

12   9D8B6DA8EC514B8

    Eugene Lee

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28